Comparing WTP values of different types of QALY gain elicited from the general public

Pennington, Mark; Baker, Rachel; Brouwer, Werner; Mason, Helen; Gyrd-Hansen, Dorte; Robinson, Angela; Donaldson, Cam; EuroVaQ Team

Published in:
Health Economics

DOI:
10.1002/hec.3018

Publication date:
2015

Document Version
Peer reviewed version

Comparing WTP values of different types of QALY gain elicited from the general public

Mark Pennington PhD[1], Rachel Baker PhD[2], Werner Brouwer PhD[3], Helen Mason PhD[2], Dorte Gyrd Hansen PhD[4], Angela Robinson MSc[5], Cam Donaldson PhD[2] and the EuroVaQ Team*

1. London School of Hygiene & Tropical Medicine
2. Yunus Centre for Social Business & Health, Glasgow Caledonian University
3. Institute of Health Policy & Management, Erasmus University Rotterdam, Netherlands
4. Institute of Public Health - Health Economics Unit, University of Southern Denmark & Danish Institute for Health Services Research
5. School of Medicine, Health Policy & Practice, University of East Anglia

*The EuroVaQ Team comprises: Cam Donaldson, Rachel Baker, Helen Mason (Glasgow Caledonian University & Newcastle University); Mark Pennington (London School of Hygiene & Tropical Medicine and Newcastle University); Sue Bell, Michael Jones-Lee, John Wildman (Newcastle University); Emily Lancsar (Monash University, Melbourne); Angela Robinson, Philomena Bacon (University of East Anglia); Jan Abel Olsen (University of Tromso), Dorte Gyrd-Hansen, Trine Kjaer, Mickael Beck, Jytte Seested Nielsen (University of Southern Denmark); Ulf Persson, Annika Bergman (Swedish Institute of Health Economics, Lund); Christel Protière, Jean Paul Moatti (INSERM, U-912, Université Aix-Marseille); Stephane Luchini (GREQAM, Centre National de la Recherche Scientifique, Marseille); Jose Luis Pinto Prades (Glasgow Caledonian University & Pablo de Olavide University, Seville); Awad Mataria, Rana Khatib, Yara Jaralla (Birzeit University, OPT); Werner Brouwer, Job van Exel (Erasmus University, Rotterdam); Roman Topór-Madry, Adam Kozierkiewicz, Darek Poznanski, Ewa Kocot (Jagiellonian University Medical College, Krakow, Poland); László Gulácsi, Márta Péntek (Corvinus University Budapest); Samer Kharroubi, Andrea Manca (University of York) ; Phil Shackley (University of Sheffield)

Address for correspondence:

Mark Pennington

Department of Health Services Research and Policy

London School of Hygiene and Tropical Medicine

15-17 Tavistock Place

London WC1H 9SH

Email: mark.pennington@lshtm.ac.uk

Tel +44 (0)20 7927 2780

Fax +44 (0)20 7927 2701

Key words: contingent valuation, willingness-to-pay, EuroVaQ

Word count: 5040

Financial support: this work was supported by a grant from the European Union Sixth Framework Programme SP5A-CT-2007-044172. The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

Key words: contingent valuation, willingness-to-pay, EuroVaQ

Abstract

Background: The appropriate thresholds for decisions on the cost-effectiveness of medical interventions remain controversial, especially in 'end of life' situations. Evidence of the values placed on different types of health gain by the general public is limited.

Methods: Across nine European countries, 17,657 people were presented with different hypothetical health scenarios each involving a gain of one quality adjusted life year (QALY) and asked about their willingness to pay (WTP) for that gain. The questions included quality of life (QoL) enhancing and life extending health gains, and a scenario where respondents faced imminent, premature death.

Results: Mean WTP values for a one-QALY gain comprised of QoL improvements were modest (PPP$11,000). When comparing QALY gains obtained in the near future the valuation of life extension exceeded the valuation of QoL enhancing gains (mean WTP PPP$19,000 for a scenario in which a coma is avoided). Mean WTP values were higher still when respondents faced imminent, premature death (PPP$29,000).

Conclusions: Evidence from the largest survey on the value of health gains by the general public indicated a higher value for life extending gains compared to QoL enhancing gains. A further modest premium may be indicated for life extension when facing imminent, premature death.

1. Introduction

Implicit or explicit recognition of the need to ration care has contributed to the formation of institutions in a number of countries (e.g. National Institute of Health and Clinical Excellence (NICE) in the UK) whose duties include evaluating whether new health technologies should be provided by publicly funded health care systems. Underpinning these decisions is some measure of the benefit of the health technology evaluated; typically the Quality Adjusted Life Year (QALY), a measure which combines the impact of health changes on both health related quality of life (QoL) and life-years. Quantifying health gains in this manner allows decision making across a range of health programmes that maximises health gain (as assessed in QALYs) for a given health care budget. It also provides a measure to examine whether the distribution of health gains across society is equitable.(1)

Decision making on the basis of a cost per QALY gained from a health technology requires stipulation of a threshold above which technologies are considered poor value for money. Methods to determine this threshold are contested. Conventional welfare economics would equate a health gain with the mean of individual willingness to pay (WTP) for the gain. Some researchers have advocated determining the general public's WTP for a QALY as a means of setting an appropriate threshold.(2,3) Others have criticised WTP as an appropriate measure of the value of health gains, arguing instead that economists should seek to determine the shadow price of health gains given a health care budget determined through a political process.(4,5) Recently, it has been suggested that a combination of these two approaches is required.(6) Both approaches rest on an assumption that QALYs appropriately quantify health gains.

Studies have shown that QALYs gained in different circumstances are valued differently. Evidence suggests clinicians place a higher value on life extending QALYs compared to life enhancing QALYs.(7) Evidence based on modelling from existing general public studies also suggests that values placed on a QALY may vary across QoL enhancing and life extending health gains.(8) The value of life extension close to death in HIV patients may approach four times the standard estimates of the monetary value of treatment effects.(9) It is unclear whether we can generalise values from a very specific clinical situation (such as the HIV patient above) to other scenarios. Debates about whether the value patients place on one QALY is generic or varies dependent on context have therefore persisted for many years.(10) Recently, NICE has indicated that a higher weight will be applied, in certain circumstances, to life-extending QALYs gained by patients who have a short life expectancy.(11) Policies initiated in Denmark, Belgium, Canada, Ireland and New Zealand (12-16) also suggest that politicians in many countries prioritise life extensions when facing premature death over other health gains.

This paper reports selected results from the EuroVaQ study (European value of a QALY). EuroVaQ was funded by the European Commission to test different ways of deriving a monetary value of a QALY through surveys of the general public.(17) Two novel internet delivered surveys were developed and completed by 39,459 respondents across ten countries; Denmark, France, Hungary, Netherlands, Norway, Palestine, Poland, Spain, Sweden, UK. Two main elicitation approaches were used, the 'direct' and the 'chained' approach, of which only the former is relevant to this paper. Both approaches involved eliciting respondents maximum WTP for a health gain using an open-ended payment system. In the 'chained' approach respondents completed a standard gamble or time trade-off exercise to value a health state.(18) They were then presented with varying risks or durations of that health state such that the overall health gain was 0.05 or 0.1 QALYs. In

contrast, the 'direct' approach comprised scenarios offering health gains of predominantly 1 QALY. We report results from five questions from the direct questionnaire which were designed to allow comparison of WTP for different *types* of QALY (QALYs comprised of different Qol gains and life extensions). At the time of the study the Principal Investigator was based at Newcastle University in England and the project met the requirements of that University's Faculty of Medicine Ethics committee.

2. Methods

The 'direct' survey avoided reference to specific health states and aimed to minimise cognitively challenging valuations. Conducting the survey over the internet increased the potential sample size and allowed questions to be customized to individual respondent characteristics (such as age and life expectancy), whilst maintaining the same nominal health gain. Respondents were presented with health gains of predominantly one QALY. This contrasts with the more common approach where respondents are presented with a very small QALY gain, usually in a risk-based format, which requires multiplying up appropriately. This direct approach was developed as a response to the methodological limitations of previous research which include limited discrimination between small risks, and unwillingness to state a valuation for certain but small health gains.(19) A full description of the methods is available in the final report at

http://research.ncl.ac.uk/eurovaq/EuroVaQ_Final_Publishable_Report_and_Appendices.pdf .(17)

2.1 Health gains valued

The five WTP questions which are the focus of this paper are highlighted in bold in Table I. Each presented a 1 QALY gain. With the exception of question C each question was presented as a health gain which would arise in a year's time. Two of the five questions (A and B) offered an improvement in QoL, a gain of 0.25 QALYs over four years (0.25*4) and a gain of 0.1 QALYs over ten years (0.1*10) respectively. The remaining three questions offered an extension of (conscious) life of one QALY. Question C (LEend) described an extension of one QALY at the end of respondents' stated life expectancy; the other two offered an extension of conscious life, but in scenarios in which the gain occurred now rather than in the future. These latter two questions were specifically designed to aid comparison with questions offering a gain in QoL by ensuring the health gains commenced at the same point in time. Question E (terminal) placed respondents in a scenario of a terminal illness. Question D (coma) presented respondents with a scenario in which they were facing a period in a coma on the assumption that time spent in a coma is equivalent to a shortening of life by that amount of time. The coma scenario was perhaps more contrived than the other questions. However, it was intended to elicit a value for a gain in longevity occurring in the near future and hence more comparable to the remaining questions than question C (LEend). The five questions, along with the remaing questions in the direct questionnaire, are summarised in Table 1A of the appendix.

There is evidence in the literature of higher valuation on life extensions compared to health gains arising from quality of life improvements.(7,8) However, the QALY gained in question C

(LEend) arises some time in the future, and we would expect values to be discounted on the assumption that individuals prefer health gains sooner than later.(20) Consequently, we hypothesised that questions D (coma) and E (terminal) would generate higher values than the remaining questions. We expected question E (terminal) to generate the highest values for two reasons. Firstly, we might expect respondents to place a higher value on an additional QALY when facing imminent death. Secondly, participants placed in a scenario of imminent death would be expected to display diminished marginal utility of income.

2.2 Question framing

First, respondents were introduced to the notion of measuring health on a visual analogue scale presented as a "health thermometer" ranging from death (zero) to full health (100). Generic health state descriptions were used to indicate different levels of health on the scale. These consisted of three EuroQoL EQ-5D-3L health states and approximate numerical valuations derived from UK survey values.(21) Respondents were then asked to estimate their remaining life expectancy and rate their current health between 0 and 100 on the 'thermometer'. From their answers to these questions and their age, subsequent scenarios for WTP questions were presented graphically and tailored to each respondent's own characteristics. The electronic questionnaire is accessible online at http://research.ncl.ac.uk/eurovaq/questionnaires.html.(22)

A screen shot of question A (0.25*4) is displayed in Figure 1. In this example a 52 year old who expects to live to 80 years of age has rated her health at 90 points. She is presented with the prospect of an illness which results in a loss of 'health' of 25 points (i.e. dropping to 65 on the thermometer) commencing in one year and lasting for four years. (We assume that a 25 point loss in health, over four years and on a 0-100 scale where 0=death and

100=full health, amounts to a one QALY loss). The respondent was then asked if she would pay anything to avoid the health loss. Payments were 'out-of-pocket', and respondents were encouraged to consider what they would be able to pay after thinking about their current income and savings and the possibility of a loan. If the respondent answered 'yes', a series of payment cards would appear at random on the screen and the respondent would be asked to sort them (via 'clicking-and-dragging' using the computer mouse) into amounts she definitely would pay, amounts she definitely would not pay and amounts about which she was unsure. The maximum card value the respondent definitely would pay and the minimum she definitely would not pay were then summarised, and the respondent asked to state her maximum WTP in an open ended response constrained by the summarised range. Nineteen cards ranging from £10 to £300,000 were converted into the respondents' local currencies (with currencies converted at 2008 purchasing power parity rates). If the respondent declined to pay to avoid the health state she was asked to indicate reason(s) from a set of pre-coded responses or using a free text option.

The remaining four questions reported here were presented in a similar manner. Question C (LEend) offered respondents the opportunity to buy an increase in life-years equivalent to one QALY. The duration and timing of the gain varied according to respondents' reported health and age at which they expected to die. Hence a respondent reporting a health of 60 and expecting to die at age 75 would be offered 20 months of additional life commencing at age 75. In question E (terminal) respondents were presented with the opportunity to delay imminent, premature death from a life threatening disease. The duration of the delay was adjusted according to the respondent's own health so that the health gain amounted to one QALY. A screen shot of this question is displayed in Figure 2. Finally, question D (coma)

presented respondents with a scenario in which the individual would slip into a coma for an amount of time equal to one QALY, whereby the QALY would be lost imminently. Respondents were asked to assume that following the coma they would return to their current state of health and 'pick up where you left off'.

Respondents were excluded from question A (0.25*4) or B (0.1*10) if their resulting QoL would have fallen below 10 points. Respondents indicating that they expected to live less than six years were directed to another version of the questionnaire which included only questions offering health gains of short duration, and excluded question E (terminal). They are not reported here. Consequently, question E represented a scenario in which death would occur at least 6 years before respondents expected to die.

2.3 Question ordering

Table I illustrates the ordering of the relevant questions within each version of the questionnaire. Version 1 contained question A (0.25*4) as the first question. In version 2 questions C (LEend) and D (coma) were randomised to first or second, with question E (terminal) presented last. Version 3 contained question B (0.1*10) as the first question. Version 4 contained all five questions described above with questions A (0.25*4) and B (0.1*10) appearing first and second in random order. They were followed by questions C (LEend) and D (coma); also in random order. Question E (terminal) was the final question.

2.4 Questionnaire Development and recruitment of subjects

Question feasibility and validity were examined using focus groups in which pilot respondents (n>50 from four European countries) provided detailed feedback prior to development of the electronic survey. The pilots took place in computer suites with support

and observation from researchers and respondents' comments were audio-recorded. Respondents were asked to comment on the credibility and clarity of the scenarios presented and their comments were used to refine and improve the presentation of questions. A prototype of the online survey was then tested in five countries and refined prior to launching the survey.

We aimed to recruit a sample of 1500 respondents per country, and for these to be representative by age and gender, socio-economic status and region. Respondents were recruited from internet panels managed by *Survey Sampling International* (SSI; http://www.surveysampling.com/). In Palestine, where it was not possible to access an internet panel, a survey invitation was emailed to a list of contacts. The survey was launched on 23 November 2009 and closed on 28 February 2010. Due to recruitment methods (inviting a large number of panel members to participate by email and closing when a sufficient number and types of respondents were achieved) standard response rates cannot be calculated. Drop-out rates for those commencing the survey were 32-48% across the nine European countries. These are higher than the typical rates for surveys from this company, which may reflect the complexity of the questions asked or the length of the questionnaire. Respondents were allocated to one of four questionnaire versions randomly until quotas for socio-demographic characteristics were achieved for each version in an attempt to achieve representativeness both within versions and across the entire sample. Respondents from demographic strata for which a sufficient number of responses had been collected across all surveys were screened out. Respondent characteristics are tabulated in the appendix (Table 2A). With the exception of Palestine, the final sample was broadly representative but with significant under representation of elderly females and lower-

income groups in some European countries. The 300 responses received from Palestine were clearly unrepresentative and consequently we excluded them.

2.5 Data analysis

The use of open ended questions allowed determination of mean and median values using the raw data avoiding the need to fit a regression model. We report mean and median WTP for a QALY derived from each question. As is conventional in WTP studies, we excluded 'protest respondents' who were not willing to pay anything with the sole reason given that 'the government should pay'.(23) The other reasons for not being willing to pay were more difficult to classify as 'protestors' or 'true zeros'. All responses other than 'government should pay' were interpreted as indicating a 'true' value of zero for the health gain in the base case analysis. Respondents who indicated *both* a protest and other reason(s) for not wishing to pay were similarly assumed to value the health gain at zero.

Means and medians are reported in $USPPP, in order to express monetary values elicited across countries in comparable currency. Country specific data and comparisons are available in the main report.(17) Confidence intervals around means within questionnaire versions were estimated using a two-stage bootstrapping procedure to allow for the hierarchical nature of the data across countries; we report bias corrected and accelerated confidence intervals.(24) We calculated 95% confidence intervals for the mean difference in WTP responses for selected question pairs using a two-stage bootstrapping routine; we also report p values for the paired t test for these question pairs. The latter test should be interpreted with caution given the extremely skewed distribution of the data. We undertook three sensitivity analyses. First, as undertaken previously in the value of life literature (25), in order to reduce the effect on means of extreme (arguably implausible) upper-end

responses, we 'trimmed' means and medians by excluding the top 1% WTP responses in each country. Second, we report means and medians after excluding *all* respondents who chose not to pay for health gains. Third, to investigate the impact of ordering effects we report means and medians after including only questions answered first in the questionnaire.

We tested whether responses differed by questionnaire version using a Kruskal-Wallis test. We investigated whether differences arose due to differences in the likelihood of expressing a positive WTP or the magnitude of positive responses differed across questionnaire versions using, respectively, a Logit regression of WTP responses dichotomised as zero and non-zero (positive) values, and an ordinary least squares (OLS) regression of the log transformed positive WTP responses allowing for clustering at country level. This regression analysis was also used to investigate the impact of respondent characteristics (including income, education, social class, household size, age and sex) on the likelihood of a positive WTP response and the magnitude of positive responses. Details of this analysis are given in the appendix.

All analysis was undertaken in Stata version 12.

3. Results

Around 8000 respondents answered each of the five questions (A to E) across the four questionnaire versions. Those electing not to pay for health gains ranged from 20% for question A (0.25*4) to 45% for question C (LEend) (Table II). A consistently small number (6-8%) of respondents were labelled protestors. The reasons for refusing to pay are tabulated in the appendix (table 3A). Table III reports results by questionnaire version for questions A

to E. Clearly the WTP response data was highly skewed. Across the four versions mean values for a one QALY gain, (rounded) ranged from $8,000 to $31,000. In the responses to version 4 of the questionnaire confidence intervals around mean values for questions A (0.25*4), D (coma) and E (terminal) did not overlap indicating a significant difference. However, the same questions in questionnaire versions 1 and 2 generated confidence intervals that were not entirely distinct. The median values for question C (LEend) were particularly low, due to the very large number of zero responses for this question. For the remaining four questions, median values (rounded) ranged from $1,100 to $2,300.

Table IV reports mean and median values after aggregation of responses for questions A to E across questionnaires. For completeness, aggregated values for the remaining questions (F to J) are reported in the appendix (Table 4A). The mean value for question D (coma) was approximately 75% higher than the mean for question A (25*4), and the mean value for question E (terminal) was about 50% higher than that for question D. The relative difference in medians across the three questions was smaller. Table IV also reports the sensitivity analyses. After trimming the top 1% of WTP-per-QALY values in each country, means were reduced by one third relative to the base case but proportional differences remained broadly the same. After excluding respondents who elected not to pay for the health gain, mean and median responses increased for all questions. The impact was greater for medians, especially for the longevity gains. The proportional difference in median values across questions A (25*4), D (coma) and E (terminal) for respondents electing to pay approached the magnitudes observed for mean values in the base case. The impact of including only questions answered first by respondents is modest for questions A (25*4), B (0.1*10) and C (LEend). Mean and median values for question D (coma) are higher.

Table V reports means and 95% confidence intervals of the difference in WTP responses across selected questions for respondents answering each respective pair of questions. The difference in means across questions A (25*4), D (coma) and E (terminal) were statistically significant. The paired t tests confirm the observations of significant differences between questions A, D and E.

The results of the regression analyses of WTP values on respondent characteristics for each of the five questions are reported in the appendix (Tables 5A and 6A). Respondents in the Netherlands, UK and France were least likely to opt to pay for health gains, and respondents in Poland and Hungary were most likely. Respondents were more likely to pay if their education level, social class or income was higher, but the trend was weak for education and social class. A greater propensity to pay amongst older respondents was observed for questions A (0.25*4), B (0.1*10) and C (LEend), but not D (Coma) and E (Terminal). Men were less likely to elect to pay, as were healthier respondents. The coefficients on the terms for question order and version 4 indicated a lower propensity to pay for health gains in questions B, C and D when these question appeared later in the question sequence.

After adjusting for respondent characteristics mean positive WTP values were highest for respondents from Spain, Denmark and Norway and lowest for respondents from the France, Hungary and Poland. Income, education and social class showed the expected relationship with WTP responses although the trend was weaker for social class. For questions A (0.25*4) and B (0.1*10) older age was associated with higher WTP. In contrast, questions D (coma) and E (terminal) showed the opposite trend with respect to age. Respondents reporting higher health reported higher WTP values for questions A (0.25*4), D (coma) and

E (terminal). The impact of gender was small, but men gave higher values. The impact of questionnaire version (reported as variable 'version 4') was not significant.

Kruskal-Wallis tests confirmed the observation of differences in the distribution of responses to questions C (LEend) and D (coma) across versions 2 and 4 ($p = 0.03$ and $p = 0.01$ respectively) but not question E (terminal) ($p = 0.38$). Kruskal-Wallis tests also confirmed a difference between versions 3 and 4 in responses to question B (0.1*10) ($p = 0.004$), but not question A (0.25*4).

4. Discussion

From the results, a mean value ranging from $10,000 to $30,000 can be placed on one extra QALY estimated in scenarios involving certainty. Respondents differentiate between gains in life years in the near future and gains in QoL reflecting patterns suggested in earlier work (7,8). The advantage of our study is that values are elicited directly from large samples of the public as opposed to being indirectly modelled (8) or elicited from physicians.(7) Comparing these figures to the relevant stated-preference literature, the magnitude of responses reflect those of other studies in which values of QALY gains have likewise been elicited under conditions of certainty.(25-27) Studies in which respondents pay to avoid a small risk of a health loss typically generate higher values, such as that by Haninger and Hammitt which used risk of food poisoning in questions specifically designed to elicit a WTP per QALY and report values between $152,000 and $5,587,000 per QALY.(28) Similarly, studies which have used the Value of a Statistical Life as a basis for calculating the value of a QALY also generally result in higher WTP per QALY values than studies which have been specifically designed for this purpose.(18,29)

Mean values for QoL gains are invariant to the shape of the gain in terms of QoL and duration. The mean value for the life extending gain (LEend) is similar to the 0.25*4 gain. However, we would expect gains offered in the more distant future to be discounted which limits comparability with the QoL gains.(30) Assuming that time spent in a coma represents a shortening of life rather than a reduction in QoL, mean values indicate that QALYs comprised of gains in life years are valued around 75% higher than those comprised of gains in QoL. A further premium is evident for health gains in an end-of-life (terminal illness) scenario, the magnitude of which depends upon the comparator. Compared to gains in QoL, extension of life close to death is valued over 150% higher; compared to gains in longevity occurring at the same point in time (coma) the premium is 50%. Evidence that the public places a greater value on life extending QALYs compared to quality of life enhancing QALYs would suggest that the use of QALYs in cost-effectiveness analysis without additional weighting is inconsistent with welfare economics.(31)

Inference from mean values is consistent with welfare theory.(32) Nevertheless, the data is highly skewed. Hence, it is reassuring to observe the same relative differences in mean values across the questions after excluding the top 1% of values. Evidence from median values that gains in longevity occurring now are valued higher than gains in QoL is weaker. However, median values are strongly influenced by large proportions of observations at zero as indicated by the very low value for question C (LEend). After excluding respondents who declined to pay (and were assigned a WTP of zero) the relative difference in median values across the five questions approaches the magnitude observed with the means.

The size of the survey allowed us to identify differences in valuations of different types of health gain despite the highly skewed distribution of responses. We used a non-parametric,

two-stage bootstrapping routine to quantify sampling uncertainty in the data. This respects the hierarchical nature of the sampling frame and the distribution of the data. Achieving a large sample size required aggregation of data across countries, as well as age groups and other characteristics. It is likely that the relative differences in WTP values across questions vary by subgroup. Regression of log WTP values on patient characteristics demonstrates the expected relationship with household income, education and social class.

By offering all respondents the same sized health gain our approach sidesteps the issue of whether to calculate an average of each respondent's WTP per QALY or to calculate average values for WTP and health state valuations prior to deriving a WTP per QALY value (the so-called mean-of-ratios or ratios-of-means problem).(33,34) It may be less cognitively demanding than the traditional approach to the Value of a Statistical Life where respondents are asked to value avoiding a small risk of a health loss.(35) The graphical displays were designed to further aid cognition. The electronic medium allowed us to present health gains in a personally relevant manner whilst maintaining the same nominal health gain across respondents. These aspects may have enhanced the validity of responses. However, the relatively large proportion of respondents electing not to pay for health gains might indicate that some respondents did not engage with the questionnaire. In this respect valuations obtained from an internet based survey may have less validity than a face-to-face questionnaire. Although active efforts were made to collect a representative sample we were constrained by the use of internet survey panels and struggled to recruit elderly women in some countries. It is possible that the exclusion of respondents who did not complete the questionnaire has introduced bias into estimated means and medians. However, it seems less likely that the differences we find in the magnitude of WTP per QALY

between questions are attributable to this. A further criticism that might be levelled at this approach is that valuations of large health gains are depressed by budget constraints.(36) Finally, we have assumed that the combination of 'VAS type' health ratings and duration, chosen to facilitate understanding of the questions, can be used to represent QALYs. This approach has been previously used in the derivation of QALY weights.(37)

Evidence from across the four versions indicates a higher proportion of respondents electing not to pay when questions appeared later in the questionnaire, which might be attributable to respondent fatigue. This trend has been previously observed (38), and may explain the rather surprising observation that fewer respondents elected to pay for the health gain in question D (coma) when compared to questions A (0.25*4) and B (0.1*10), and fewer still elected to pay in question E (terminal). In this respect our estimate of the difference in values between questions A, D and E in the base case analysis is likely to be conservative. The relative differences are probably larger, as indicated by the sensitivity analysis which excluded questions not appearing first in the questionnaire version seen by each respondent.

Reasons given by respondents for declining to pay for the health gain in the terminal illness scenario suggest that, despite the wording of the question, many of the respondents assumed the extra time would be spent in poor health. Although life extensions were described to respondents at their own current QoL as indicated on the thermometer, perceptions of QoL may well have been affected by the fact that people in our hypothetical question know they are going to die. Hence, all notions of hope are excluded. In the real world setting people still have a hope that they may live for a long time.(39) In this sense our form of question would inevitably lead to conservative valuations.

One interpretation of these results is that they are consistent with findings that the general public place a higher social value on health gains achieved from a lower or more severe health state.(,40,41) This need not surprise, since individually and socially focussed approaches do not necessarily differ in this regard, although aspects like solidarity and equity enter only in the latter.(42,43) Whilst we have interpreted question D (coma) as offering an extension of life, it may also be interpreted as a gain in QoL from a severe state (zero), and higher valuations may reflect this interpretation. One could further argue that question E (terminal) offers a health gain in a more severe scenario, although responses are more influenced in this scenario by the 'dead anyway' effect,(44) thus lowering many respondents' perceptions of their marginal utility of income. Whilst such consistency is clearly possible the data are not entirely congruent with this argument. For example, there is no difference in mean values across questions A (0.25*4) and B (0.1*10) despite one offering a health gain from a lower initial health than the other. Of course, this could be because each of these is perceived as not being greatly different from each other in terms of severity.

Clearly some caution is needed in interpreting the differences in values observed across the questions. Although respondents were reassured that they would be able to pick up their life as before on emerging from the coma, respondents may not have found this plausible and the resulting values may be inflated by concerns regarding the impact of the coma on relationships and employment.

Caution is also needed when interpreting the premium on health gains when facing imminent death. Whilst the premium is large when compared to gains in QoL only, a contaminating factor in this comparison is the different types of health gain offered: QoL

gains in one scenario with life extension in the other. Question D (coma) was designed to overcome this, and the problem of life extension being interpreted by respondents as adding on time at the end of life. When comparing the mean for question D (coma) with that for question E (terminal), the 'premium' is reduced to 50%.

One may also question the appropriateness of applying the valuations expressed in question E (terminal) as a guide for allocating public (or private) insurance resources. A core concept in economics is that all investments are associated with opportunity costs since investment of resources in a given project will initiate reduced possibilities (and thus loss of utility) elsewhere.(45) For the person who is facing death this rule may not apply, since money may be more or less worthless to the individual and the notion of opportunity costs nonsensical, although this is tempered somewhat by a desire to leave a legacy. Hence, it is natural for the person facing the prospect of imminent death to express high valuations (in terms of WTP) for life extensions, both because such life extensions are highly valued (either because doing something initiates hope, or because one values the extra time spent with family and friends) *and* because opportunity costs are low or nonexistent. It is therefore perfectly understandable and may indeed be utility maximising behaviour for such individuals to spend large sums of money out-of-pocket on interventions, which in reality offer only a chance of a minor life extension. However, if the health care intervention is financed by a common pool of funds (be this public or private) it is the general public (who are not facing death in the near future) who are to finance the health care intervention, and for whom the opportunity cost of spending additional dollars on terminal illness programmes may well be high in terms of other programmes forgone.

5. Conclusions

Results from the largest ever survey of the general public support previous findings of modest valuations of health gains elicited under certainty. This paper has demonstrated some of the complexities associated with trying to elicit values for different types of one-QALY gain, especially for different forms of life-extending QALYs: eliciting preferences for health amongst patients who are facing death is not straightforward, as it conflates issues of QoL versus length of life as well as whether such gains happen in a terminal illness as opposed to another scenario.

Findings in this study support previous evidence that life extending health gains are valued more highly than QoL gains. Evidence for a premium on end-of-life (terminal illness) QALYs depends on how the results are interpreted. On one hand, compared with typical contexts for gaining QALYs, such a premium would appear to exist and to reflect judgments of major health technology assessment bodies. On the other hand, relative to another life extending scenario designed specifically to aid comparison with terminal illness, the premium is much smaller. Given the likelihood that utility of income falls in a terminal illness scenario, this comparison would indicate that any premium is modest.

6. Acknowledgements

All authors declare: no financial relationships with any organisations that might have an interest in the submitted work, no other relationships or activities that could appear to have influenced the submitted work. The research leading to these results received funding from the European Community's Sixth Framework Programme under grant agreement n° 044172 – the EuroVaQ project. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

7. References

1. Schwappach DLB. Resource allocation, social values and the QALY: a review of the debate and empirical evidence. 2002. *Health Expectations* **5**: 210.
2. Hirth RA, Chernew ME, Miller E, Fendrick MA, Weissert WG. 2000. Willingness to pay for a quality-adjusted life year: in search of a standard. *Medical Decision Making* **20**: 332.
3. Weinstein, MC. 2008. How Much Are Americans Willing to Pay for a Quality-Adjusted Life Year? *Medical Care* **46**(4):343
4. McCabe C, Claxton K, Culyer A. 2008. The NICE cost-effectiveness threshold: what it is and what that means. *Pharmacoeconomics* **26**: 733–44.
5. Culyer AJ, McCabe CJ, Briggs AH, Claxton K, Buxton M, Akehurst R et al. 2007. Searching for a threshold not setting one: the role of the National Institute for Health and Clinical Excellence. *Journal of Health Services Research and Policy* **12**: 56–8.
6. Baker R, Chilton S, Donaldson C, Jones-Lee M, Lancsar E, Mason H et al. Searchers vs surveyors in estimating the monetary value of a QALY: resolving a nasty dilemma for NICE. 2011. *Health Economics, Policy and Law* **6**: 435
7. Kozminski MA, Neumann PJ, Nadler ES, Jankovic A, Ubel PA. 2011. How long and how well: oncologists' attitudes toward the relative value of life-prolonging v quality of life-enhancing treatments. *Medical Decision Making* **31**: 380-385.
8. Mason H, Jones-Lee MW, Donaldson C. 2009. Modelling the monetary value of a QALY: a new approach based on UK data. *Health Economics* **18**: 933.

9. Philipson TJ, Becker G, Goldman D, Murphy KM. 2010. Terminal care and the value of life near its end. Working Paper 15649, National Bureau of Economic Research, Cambridge, Mass. http://www.nber.org/papers/w15649 [25 August 2011].

10. Donaldson C, Atkinson A, Bond J, Wright I. 1988. Should QALYs be programme-specific? *Journal of Health Economics* **7**:239-257.

11. National Institute for Health and Clinical Excellence. 2009. Appraising life-extending, end of life treatments. http://www.nice.org.uk/media/88A/F2/SupplementaryAdviceTACEoL.pdf [10 October 2012].

12. Jonsson B, Willking N. 2007. Market access for cancer drugs and the role of health economics. *Annals of Oncology* **18**: 55-66.

13. Corens D. 2007. Belgium. In *Health systems in transition*. World Health Organisation and European Observatory of Health Systems and Policies: Copenhagen.

14. Berry SR, Hubay S, Soibelman H, Martin DK. 2007. The effect of priority setting decisions for new cancer drugs on medical oncologists' practice in Ontario: a qualitative study. *BMC Health Services Research* **7**: 193.

15. Barry M, Tilson L. 2007. Recent developments in pricing and reimbursement of medicines in Ireland. *Expert Review of Pharmacoeconomics & Outcomes Research* **7**: 605-611.

16. Raftery J. 2008. Paying for costly pharmaceuticals: regulation of new drugs in Australia, England and New Zealand. *Medical Journal of Australia* **188**: 26-28.

17. European Value of a Quality Adjusted Life Year, project no. SP5A-CT-2007-044172**,** 2010.

http://research.ncl.ac.uk/eurovaq/EuroVaQ_Final_Publishable_Report_and_Appendices.pdf [16 June 2010].

18. Robinson A, Gyrd-Hansen D, Bacon P, Baker R, Pennington M, Donaldson C, EuroVaQ Team. 2013. Estimating a WTP-based value of a QALY: the 'chained' approach. *Social Science & Medicine* **92**: 92-104.

19. Baker R, Bateman I, Donaldson C, Jones-Lee M, Lancsar E, Loomes G, *et al*. 2010. Weighting and valuing quality-adjusted life-years using stated preference methods: preliminary results from the Social Value of a QALY Project. *Health Technology Assessment* **14** (27).

20. Porat A, Tabbach AD. 2010. Risk of Death. University of Chicago Law and Economics Olin Working Paper No. 531. http://www.law.uchicago.edu/files/file/531-ap-death.pdf [22 September 2010]

21. Dolan P, Gudex C, Kind P, Williams AA. 1999. Social Tariff for EuroQoL: Results from a UK General Population Survey. CHE Discussion Paper 138, York: Centre for Health Economics, University of York, UK.

22. Online questionnaire available at http://research.ncl.ac.uk/eurovaq/questionnaires.html

23. Olsen JA, Donaldson C. 1998. Helicopters, hearts and hips: Using willingness to pay to set priorities for public sector health care programmes. *Social Science and Medicine* **46**:1-12.

24. Ng ES-W, Grieve R, Carpenter JR. 2012. Two stage non-parametric bootstrap sampling with shrinkage correction for clustered data. *Stata Journal*; [in press].

25. Carthy T, Chilton S, Covey J, Hopkins L, Jones-Lee M, Loomes G et al. 1999. On the Contingent Valuation of Safety and the Safety of Contingent Valuation: Part 2—The CV/SG "Chained" Approach. *Journal of Risk and Uncertainty* **17**(3): 187–213.

26. Johannesson M and Johansson P-O. 1997. Quality of life and the WTP for an increase in life expectancy at an advanced age. *Journal of Public Economics* **65**: 219-228.

27. Gyrd-Hansen D. 2003. Willingness to pay for a QALY. *Health Economics* **12**: 1049-60.

28. Bobinac A, van Exel NJA, Rutten FFH, Brouwer WBF. 2010. Willingness to pay for a quality-adjusted life-year: the individual perspective. *Value in Health* **13**: 1046-1055.

29. Haninger K, Hammitt, JK. 2011. Diminishing willingness to pay per Quality-Adjusted Life Year: valuing acute foodborne illness. *Risk Analysis* **31**(9): 1363-1380.

30. Mason H, Baker RM, Donaldson C. 2009. Willingness to Pay for a QALY: past, present and the future. *Expert Reviews of Pharmacoeconomics and Outcomes Research* **8**(6): 575-582.

31. Cairns J. 1994. Valuing future health benefits. *Health Economics* **3**:221-229.

32. Schlander M. 2010. Measures of efficiency in healthcare: QALMs about QALYs? *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen* **104**(3): 214-226.

33. Brouwer WBF, Culyer AJ, Van Exel NJA, Rutten FFH. 2008. Welfarism vs. extra-welfarism. *Journal of health economics* **27**(2): 325-338.

34. Baker RM, Bateman I, Donaldson C, Jones-Lee M, Lancsar E, Loomes G et al. 2010. Weighting and valuing quality adjusted life years: preliminary results from the social value of a QALY project *Health Technology Assessment* **14**(7).

35. Gyrd-Hansen D, Kjær T.2012. Disentangling WTP per QALY data: different analytical approaches, different answers. *Health Economics* **21**(3): 222-237.

36. Beattie J, Covey J, Dolan P, Hopkins L, Jones-Lee M, Loomes G et al. 1998. On the contingent valuation of safety and the safety of contingent valuation: Part 1 – caveat investigator. *Journal of Risk and Uncertainty* **17**(1): 5-26.

37. Baron J, Greene J. 1996. Determinants of insensitivity to quantity in valuation of public goods: Contribution, warm glow, budget constraints, availability, and prominence. *Journal of Experimental Psychology: Applied* **2**(2): 107-125.

38. Lancsar E, Wildman J, Donaldson C, Ryan M, Baker R. 2011. Deriving distributional weights for QALYs through discrete choice experiments. *Journal of Health Economics* **30**(2): 466-478.

39. Stewart JM, O'Shea E, Donaldson C, Shackley P. 2002. Do ordering effects matter in willingness-to-pay studies of health care?. *Journal of health economics* **21**(4): 585-599.

40. Snyder, CR. 1994. *The Psychology of Hope*. Free Press, New York.

41. Nord E. 2005. Concerns for the worse off: fair innings vs severity. *Social Science & Medicine*. **60**(2): 257-263.

42. Shah, KK. 2009. Severity of illness and priority setting in health care: A review of the literature. *Health policy* **93**: 77-84.

43. Bobinac A, van Exel NJA, Rutten FFH, Brouwer WBF. Valuing QALY gains applying a societal perspective. *Health Economics,* in press.

44. Bobinac A, van Exel NJA, Rutten FFH, Brouwer WBF. 2012. Inquiry into the relationship between Equity Weights and the value of a QALY. *Value in Health* 15: 1119-1126.

45. Pratt JW, Zeckhauser RJ. 1996. Willingness to Pay and the Distribution of Risk and Wealth. *Journal of Political Economy,* 747-763.

46. Palmer S, Raftery J. 1999. Economics notes: opportunity cost. *BMJ British Medical Journal* **318**(7197): 1551.

8. Tables

Table I. Questionnaire layout

<table>
<tr><th>Version 1</th><th colspan="2">Version 2</th><th colspan="2">Version 3</th><th colspan="2">Version 4</th></tr>
<tr><td>Question A (0.25*4)

25 point loss over 4 years now</td><td>Question C (LEend

extra life at end of life</td><td>Question D (coma)

extra life now</td><td colspan="2">Question B (0.1*10)

10 point loss over 10 years now</td><td>Question A (0.25*4)

25 point loss over 4 years now</td><td>Question B (0.1*10)

10 point loss over 10 years now</td></tr>
<tr><td>Question F

25 point loss over 4 years near end of life</td><td colspan="2" rowspan="2"></td><td colspan="2">Question G

10 point loss over 10 years near end of life</td><td colspan="2" rowspan="2"></td></tr>
<tr><td>Question M

10% risk of 10 point loss over 10 years now</td><td>Question K

5% risk of 25 point loss over 4 years now</td><td>Question L

10% risk 25 point loss over 4 years now</td><td>Question I

25 point loss over 1 year now</td><td>Question J

10 point loss over 1 year now</td></tr>
<tr><td>Question H

25 point loss over 4 years now with payment over 4 years</td><td colspan="2">Question E (terminal illness)

extra life now</td><td colspan="2" rowspan="2"></td><td colspan="2">Question E (terminal illness)

extra life now</td></tr>
<tr><td colspan="3"></td><td colspan="2"></td></tr>
</table>

Questions shown in adjacent boxes separated by a dashed line were presented in random order to participants.

Table II. Scenarios over which a 'value of one QALY' was elicited

Scenario description	Number of respondents across nine countries	Number of respondents electing not to pay	Number of respondents classified as protesters
Qu. A (25*4)	8511	1716 (20.2%)	541 (6.4%)
Qu. B (0.1*10)	8374	1899 (22.7%)	545 (6.5%)
Qu. C (LEend)	8955	4027 (45.0%)	656 (7.3%)
Qu. D (Coma)	8933	2588 (29.0%)	751 (8.4%)
Qu. E (terminal)	8745	2931 (33.5%)	655 (7.5%)

Table III. Value of one QALY by scenario and questionnaire version

Scenario description	n*	Number of zeros (%)	Mean WTP ($USPPP) (95% CI)	Median WTP ($USPPP)	99th percentile ($USPPP)	Maximum Value ($USPPP)
Questionnaire version 4						
Qu. A (25*4)	3,957	614 (15.5%)	10,724 (8,995 – 13,178)	1,532	133,417	2,156,605
Qu. B (0.1*10)	3,803	751 (19.7%)	11,497 (9,252 – 17,455)	1,078	127,646	5,391,513
Qu. C (LEend)	3,935	1729 (43.9%)	8,309 (6,947 – 10,363)	149	125,203	1,531,755
Qu. D (Coma)	3,889	973 (25.0%)	18,675 (15,583 – 22,456)	1,791	453,617	3,063,510
Qu. E (terminal)	3,949	1137 (28.8%)	30,563 (25,893 – 37,268)	2,293	515,969	4,636,701
Questionnaire version 2						
Qu. C (LEend)	4,147	1704 (41.1%)	12,986 (9,537 – 20,027)	191	229,978	6,670,838
Qu. D (Coma)	4,101	893 (21.8%)	19,385 (16,096 – 25,194)	2,149	284,551	5,318,588
Qu. E (terminal)	4,118	1186 (28.8%)	27,622 (23,252 – 35,569)	2,096	459,957	7,717,024
Questionnaire version 1						
Qu. A (25*4)	3,981	610 (15.3%)	10,763 (8,458 – 15,788)	1,334	126,488	4,669,587
Questionnaire version 3						
Qu. B (0.1*10)	4,002	656 (16.4%)	11,357 (9,720 – 13,619)	1,150	199,186	1,149,892

***** The numbers in this column exclude 'protesters' and a small number of respondents for whom the survey failed to record their WTP response.

Table IV. Value of one QALY aggregated across questionnaire versions with and without trimming at 1% and after excluding respondents electing not to pay

Scenario	WTP values aggregated across questionnaire versions (base case)			Top 1% of positive values excluded (by country)			WTP values for respondents agreeing to pay			WTP values derived from responses to respondents' first question		
	n	WTP ($USPPP)		n	WTP ($USPPP)		n	WTP ($USPPP)		n	WTP ($USPPP)	
		Mean	Median		Mean	Median		Mean	Median		Mean	Median
Qu. A (25*4)	7,938	10,744	1,450	7,867	6,992	1,433	6,763	12,610	2,163	5,974	10,165	1,379
Qu. B (0.1*10)	7,805	11,425	1,138	7,737	7,093	1,098	6,451	13,823	2,070	5,904	12,576	1,150
Qu. C (LEend)	8,082	10,709	160	8,031	6,077	155	4,701	18,411	1,791	2,053	8,683	155
Qu. D (Coma)	7,990	19,039	2,096	7,925	13,618	1,941	6,162	24,687	3,829	2,075	25,013	2,875
Qu. E (terminal)	8,067	29,062	2,196	8,005	20,928	2,149	5,792	40,477	5,392	-	-	-

Table V. Mean 'within-respondent' differences in values between questions (respondents answering more than one question)

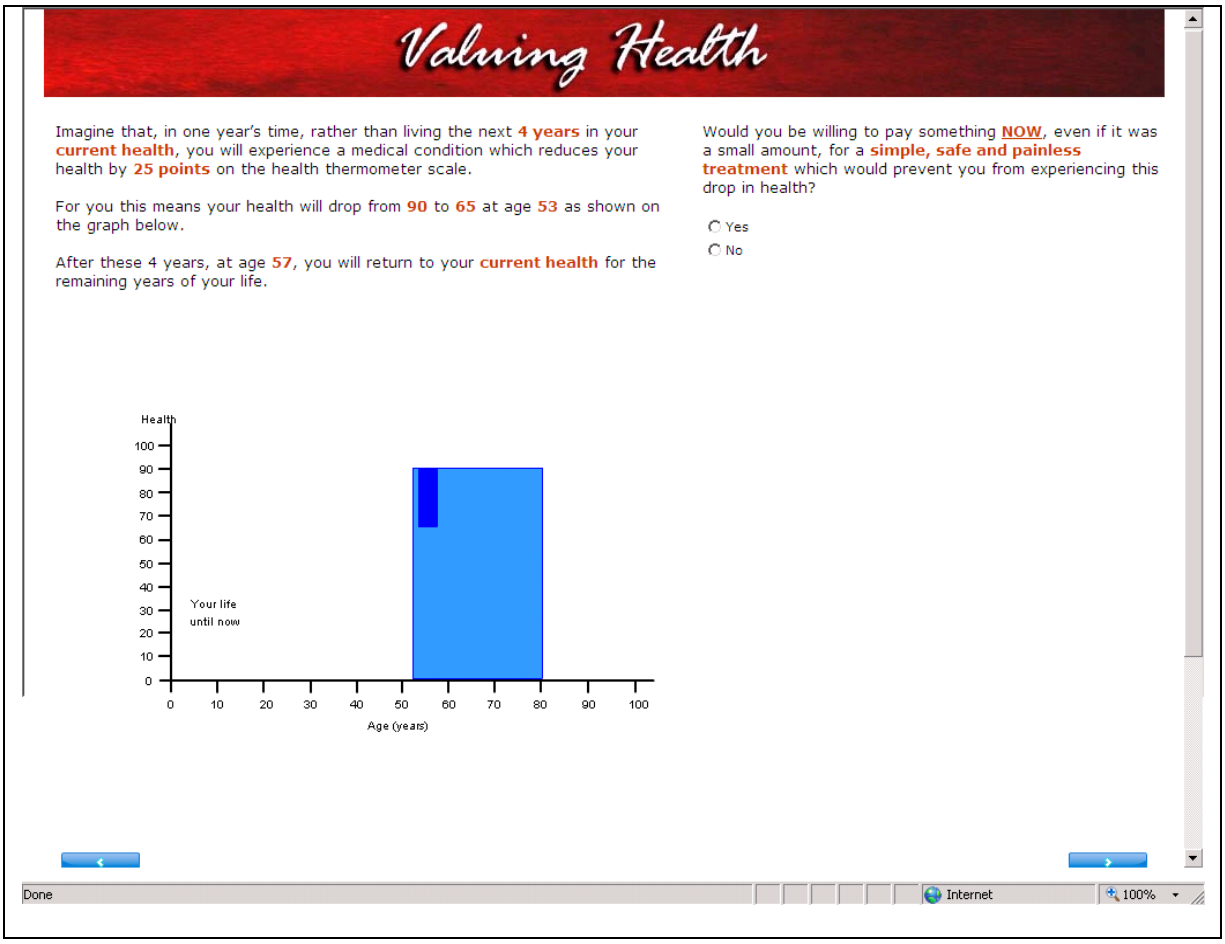
Scenarios compared	n	Mean difference ($USPPP)	95% CI	p-value for paired t test
Questionnaire version 4				
Qu. A (0.1*10) – Qu. B (0.25*4)	3,700	565	-1,183 to 4,209	0.6
Qu. C (LEend) – Qu. A (0.25*4)	3,797	-2,468	-4,412 to -737	0.008
Qu. D (coma) – Qu. A (0.25*4)	3,763	7,891	5,046 to 11,610	<0.0001
Qu. E (terminal) – Qu. D (coma)	3,790	12,218	8,048 to 18,014	<0.0001
Questionnaire version 2				
Qu. E (terminal) – Qu. D (coma)	3,975	8,831	3,773 to 16,030	0.002

9. Legends

Figure 1 **Screen shot of question A (0.25*4)**

Figure 2 **Screen shot of question E (terminal)**

10. Figures


Valuing Health
Imagine that, in one year's time, rather than living the next 4 years in your current health, you will experience a medical condition which reduces your health by 25 points on the health thermometer scale.
For you this means your health will drop from 90 to 65 at age 53 as shown on the graph below.
After these 4 years, at age 57, you will return to your current health for the remaining years of your life.
Would you be willing to pay something NOW, even if it was a small amount, for a simple, safe and painless treatment which would prevent you from experiencing this drop in health?
Yes
No
Health
100
90
80
70
60
50
40
30
20
10
0
Your life until now
0
10
20
30
40
50
60
70
80
90
100
Age (years)
Done
Internet
100%

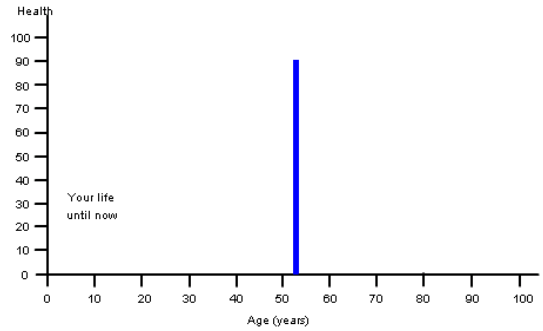
Valuing Health

Imagine that you are diagnosed with a medical condition which is **terminal** if untreated. Death from this medical condition would occur within a few months of the diagnosis if treatment is not received.

Imagine that there is a **simple, safe and painless treatment** available which would give you **13 months extra** in your **current health**.

Would you be willing to pay something **NOW** for this treatment?

- ○ Yes
- ○ No



Done Internet 100%

Appendix

Table 1A. Summary of questions forming the EuroVaQ direct questionnaire.

QALY gain	Question	Timing	Health loss/gain	Duration	Risk
1	**A**	**1 year's time**	**25 point loss**	**4 years**	**certainty**
1	**B**	**1 year's time**	**10 point loss**	**10 years**	**certainty**
1	**C**	**Near end of life**	**year of life**	**varies**	**certainty**
1	**D**	**1 year's time**	**year of life**	**varies**	**certainty**
1	**E**	**1 year's time**	**year of life**	**varies**	**certainty**
1	F	Near end of life	25 point loss	4 years	certainty
1	G	Near end of life	10 point loss	10 years	certainty
1	H*	1 year's time	25 point loss	4 years	certainty
0.25	I	1 year's time	25 point loss	1 year	certainty
0.1	J	1 year's time	10 point loss	1 year	certainty
0.1	K	1 year's time	25 point loss	4 years	10% risk
0.1	L	1 year's time	10 point loss	10 years	10% risk
0.05	M	1 year's time	25 point loss	4 years	5% risk

*In question H respondents were asked to pay for the health gain in yearly instalments

Table 2A. Respondent characteristics by country

	Netherlands	UK	France	Spain	Sweden	Norway	Denmark	Poland	Hungary	Palestine
Number of respondents	2148	1778	2088	2167	1872	1634	2234	1812	1924	300
Proportion male	49.3%	47.7%	46.7%	51.7%	44.7%	52.8%	49.5%	51.7%	50.5%	71.7%
Mean age in years (sd)	48.6 (15.0)	46.6 (16.0)	45.2 (14.6)	41.4 (13.9)	47.7 (15.7)	44.2 (15.6)	48.6 (15.7)	38.2 (13.3)	41.8 (14.2)	28.9 (9.1)
Mean stated life expectancy in years (sd)	82.3 (8.2)	81.0 (9.1)	81.2 (9.3)	83.5 (8.9)	83.5 (8.6)	83.3 (7.8)	84.1 (8.1)	78.9 (10.2)	75.9 (10.8)	64.5 (13.2)
Mean stated health - VAS scale 0 to 100 (sd)	83.1 (14.6)	80.7 (16.6)	81.5 (16.1)	83.9 (15.1)	81.8 (16.8)	84.0 (15.8)	85.1 (14.9)	82.9 (15.4)	80.8 (15.7)	77.4 (16.8)
Mean household size (sd)	2.4 (1.3)	2.5 (1.4)	2.6 (1.5)	2.8 (1.3)	2.3 (1.3)	2.5 (1.3)	2.2 (1.2)	3.2 (1.4)	2.8 (1.3)	7.5 (6.7)
Head of household*	72.7%	71.0%	66.4%	61.8%	76.8%	77.8%	76.1%	62.5%	66.8%	57.7%
Degree**	34.8%	27.8%	40.6%	36.8%	43.4%	51.5%	45.7%	53.7%	17.6%	48.3%
Respondent employed	52.1%	50.3%	50.7%	53.1%	50.3%	61.3%	50.0%	64.0%	62.0%	49.3%
Household income	72,733	95,588	47,436	62,517	59,630	106,011	112,276	38,018	19,116	3,731
Missing income	19.2%	12.0%	9.0%	15.3%	9.3%	10.3%	10.0%	14.8%	11.4%	11.0%

*Proportion of respondents who state they are the head, or the joint head of the household. **Proportion of respondents educated to degree level.

Table 3A. Reasons for not being willing to pay*

Scenario description	Could live with it	Get better anyway	Can't afford it	Govt should pay	Gain is too small	May be dead then	Health poor during gain	Wish to leave legacy	Other
Qu. A (25*4)	21%	16%	34%	45%	-	-	-	-	8%
Qu. B (0.1*10)	32%	17%	29%	43%	-	-	-	-	8%
Qu. C (LEend)	-	-	20%	30%	30%	35%	23%	-	10%
Qu. D (Coma)	23%	22%	29%	43%	-	-	-	-	9%
Qu. E (terminal)	-	-	21%	36%	-	-	48%	19%	7%

Table Key

Could live with it	It wouldn't be too bad/I could live with it.
Get better anyway	I would get better anyway, so it is not worth paying for the treatment.
Can't afford it	I do value the treatment, but I cannot afford to pay anything for it.
Govt should pay	I do value the treatment, but do not want to pay because the government should provide health care.
Gain is too small	It's not enough of a gain to be of value to me.
May be dead then	I may not live until that age, so it is not worth paying for treatment now.
Health poor during gain	I may be in poor health at that age, so it is not worth paying for treatment (life extension)/If I was going to die this would only be paying to prolong my death (terminal illness)
Wish to leave legacy	I am thinking about my family/partner – I would want to leave the money to them.

* Respondents not restricted to one answer, hence row totals exceed 100%

Table 4A: mean and median WTP per QALY for questions not reported in the main text

Scenario	n	Mean WTP (USD)	Median WTP (USD)
Gain in QoL of 1 QALY over 4 years near end of life	3,886	10,685	334
Gain in QoL of 1 QALY over 10 years near end of life	4,041	46,686	4,313
Gain in QoL of 0.25 QALYs over 1 year now	4,074	104,733	3,723
Gain in QoL of 0.1 QALYs over 1 year now	3,800	11,795	517
10% risk of loss in QOL of 1 QALY over 4 years now	4,043	119,975	3,582
5% risk of loss in QOL of 1 QALY over 4 years now	4,038	177,996	3,293
10% risk of loss in QOL of 1 QALY over 10 years now	3,750	81,740	5,319
Gain in QoL of 1 QALY over 4 years now with yearly payment	3,857	34,479	3,064

Regression of WTP responses on patient characteristics

The distribution of WTP response data was both highly skewed and included significant zero values. Consequently, regression analysis for each question was undertaken in two parts. The proportion of zero valuations was modelled as a function of respondent characteristics and questionnaire version using Logit regression. The distribution of positive response data was approximately normal after log transformation. Hence, positive WTP values were modelled using OLS regression after conversion to USD and log transformation. We found no evidence of heteroskedasticity in the log transformed data ($p = 0.9$). However, we chose to cluster on country, thus ensuring that we estimated robust standard errors.

We pre-specified the inclusion of the following covariates: age, sex, income, education, social class, household size, initial health (as reported by respondent's on the 'thermometer') and included dummy variables indicating questionnaire version and question ordering. Within each country respondents were assigned to quintiles on household income with a sixth category for respondents who failed/refused to state their income. Respondents were assigned to three levels of education based on whether they had completed compulsory education only; some further education; or had completed a degree or equivalent. Household size was analysed as the OECD coefficient which sums a score of one for the first household member; 0.5 for additional members aged over 13; and 0.3 for children under 14. Social class was assigned using ESOMAR categories which are based on job description and age on leaving education (European Society for Opinion and Marketing Research. Standard demographic segmentation. A system of international socio-economic classification of respondents to survey research 1997. Available at www.esomar.org)

An area of potential concern was the risk of endogeneity for the variable reporting respondents' health. Respondents' reported health influenced the presentation of each health gain to that respondent *via* the customisation of the questions. We tested for this with a Durbin-Wu-Hausman test and found no evidence of endogeneity ($p = 0.2$). The results of the Logit and OLS regressions for each of the questions are tabulated below.

Table 5A. Coefficients from Logit regression of WTP responses against respondent characteristics

	Coefficient	0.25*4	0.1*10	Life extension	Coma	Terminal illness
Country	Netherlands	-0.757***	-0.71***	-0.942***	-0.814***	-0.758***
	UK	-1.014***	-0.893***	-0.665***	-0.813***	-0.698***
	France	-0.464***	-0.383***	-0.516***	-0.752***	-0.849***
	Spain	-0.497***	-0.199***	-0.132***	-0.209***	-0.583***
	Sweden	-0.099***	-0.099***	-0.132***	-0.226***	-0.076***
	Denmark	-0.068***	-0.046*	-0.523***	-0.422***	-0.323***
	Poland	0.176***	0.707***	0.202***	-0.262***	-0.17***
	Hungary	0.029	0.129**	0.041	-0.176***	-0.319***
Age range	Age 18 to 25	-0.321*	-0.428**	-0.564***	0.239**	0.065
	Age 26 to 35	-0.369***	-0.387***	-0.44***	0.087	0.035
	Age 36 to 45	-0.258***	-0.261**	-0.211*	0.116	0.004
	Age 56 to 65	0.01	0.057	0.337***	0.038	0.03
	Age over 65	0.169	0.299*	0.497***	0.13	0.053
Income	Income missing	-0.577***	-0.595***	-0.294	-0.419**	-0.322*
	Income lowest	-0.601***	-0.633***	-0.291***	-0.484***	-0.317***
	Income low	-0.318**	-0.423***	-0.169	-0.149	-0.158
	Income higher	0.287*	0.004	-0.018	0.178*	0.14
	Income highest	0.197	0.227	0.206	0.326**	0.242
Education	Low education	-0.27*	-0.221**	-0.08	-0.104	-0.235**
	High education	0.135	0.094	-0.035	0.027	0.229**
Social class	Social Group A	0.233	0.313***	0.15*	0.302**	0.16*
	Social Group B	0.426***	0.234	0.058	0.121	-0.067
	Social Group C1	0.142	0.166	0.044	0.208*	0.077
	Social Group D	0.093	0.155	-0.031	0.08	-0.003
	Social Group E	-0.056	-0.059	-0.139	-0.028	-0.178
Other	Male	-0.133	-0.175*	0.088	-0.268**	-0.285***
	Household size	-0.046	-0.105	0.034	0.028	0.009
	Health (VAS)	0.004	-0.004***	-0.004***	0.011***	0.008***
	Question order	-0.011	0.323***	0.093*	-0.308***	-
	Version 4	-0.019	-0.229***	-0.14***	-0.212***	-0.011
	Constant	2.135***	2.603***	1.111***	1.010***	0.799***
	Adjusted R^2	0.057	0.064	0.048	0.049	0.041

*Coefficient with p value < 0.05, **Coefficient with p value < 0.01, ***Coefficient with p value < 0.001

Base country is Norway; base income and education category is middle; base social group is C2

For the QoL gain questions the Question order variable indicates that the 0.1*10 question appeared before the 0.25*4 question; for the Coma and life extension questions the Question order variable indicates that the life extension question appeared before the coma question

Table 6B. Coefficients from regression of log WTP responses against respondent characteristics

	Coefficient	0.25*4	0.1*10	Life extension	Coma	Terminal illness
Country	Netherlands	-0.395***	-0.117***	-0.248***	-0.441***	-0.491***
	UK	-0.343***	-0.229***	0.069*	-0.263***	-0.017
	France	-1.018***	-0.836***	-0.734***	-1.141***	-1.083***
	Spain	-0.302***	-0.03	0.174***	0.075**	0.002
	Sweden	-0.418***	-0.405***	-0.255***	-0.399***	-0.503***
	Denmark	0.13***	0.336***	0.13***	0.158***	0.147***
	Poland	-0.847***	-0.798***	-0.657***	-1.159***	-1.093***
	Hungary	-0.899***	-0.798***	-0.836***	-1.15***	-1.065***
Age range	Age 18 to 25	0.088	0.028	0.103	0.386**	0.779***
	Age 26 to 35	0.033	-0.111*	-0.088	0.224***	0.38**
	Age 36 to 45	0.018	-0.127	-0.115	0.124	0.231**
	Age 56 to 65	0.174*	0.186**	-0.002	-0.052	-0.227**
	Age over 65	0.297**	0.32*	0.132	-0.073	-0.371*
Income	Income missing	-0.115	-0.022	-0.022	-0.122	-0.114
	Income lowest	-0.502**	-0.372**	-0.511**	-0.504***	-0.521***
	Income low	-0.251*	-0.234*	-0.298*	-0.194**	-0.251**
	Income higher	0.201*	0.238**	0.16	0.172*	0.196
	Income highest	0.478**	0.587**	0.44**	0.572***	0.594***
Education	Low education	-0.253**	-0.343***	-0.189	-0.292*	-0.311**
	High education	0.18**	0.203**	0.171*	0.313***	0.368***
Social class	Social Group A	0.377**	0.266**	0.396**	0.278*	0.405**
	Social Group B	0.282**	0.261**	0.112	0.144	0.285**
	Social Group C1	0.098	0.07	0.017	0.014	0.169
	Social Group D	0.015	-0.002	0.08	-0.088	0.083
	Social Group E	-0.095	-0.14	-0.209	-0.255**	-0.16
Other	Male	0.218**	0.142*	0.182*	0.093	0.132*
	Household size	-0.126*	-0.122**	-0.189*	-0.159*	-0.181**
	Health (VAS)	0.006**	0.001	0.002	0.007**	0.009**
	Question order	0.072	-0.014	-0.436**	-0.281**	-
	Version 4	0.08	-0.007	-0.007	0.018	0.076
	Constant	7.385***	7.697***	8.009***	8.250***	8.132***
	R^2	0.123	0.118	0.109	0.150	0.157

*Coefficient with p value < 0.05, **Coefficient with p value < 0.01, ***Coefficient with p value < 0.001

Base country is Norway; base income and education category is middle; base social group is C2

For the QoL gain questions the Question order variable indicates that the 0.1*10 question appeared before the 0.25*4 question; for the Coma and life extension questions the Question order variable indicates that the life extension question appeared before the coma question